

⑬ BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ Offenlegungsschrift
⑪ DE 3834636 A1

⑤ Int. Cl. 5:
C12Q 1/68

⑰ Aktenzeichen: P 38 34 636.2
⑱ Anmeldetag: 11. 10. 88
⑲ Offenlegungstag: 19. 4. 90

DE 3834636 A1

⑦1 Anmelder:

Max-Planck-Gesellschaft zur Förderung der
Wissenschaften eV, 3400 Göttingen, DE

⑦4 Vertreter:

Vossius, V., Dipl.-Chem. Dr.rer.nat.; Tauchner, P.,
Dipl.-Chem. Dr.rer.nat.; Heunemann, D., Dipl.-Phys.
Dr.rer.nat.; Rauh, P., Dipl.-Chem. Dr.rer.nat.;
Hermann, G., Dipl.-Phys. Dr.rer.nat., Pat.-Anwälte,
8000 München

⑦2 Erfinder:

Jäckle, Herbert, Dr., 7400 Tübingen, DE; Tautz,
Diethard, Dr., 8000 München, DE

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤4 Verfahren zur Analyse von Längenpolymorphismen in DNA-Bereichen

Beschrieben wird ein Verfahren zur Analyse von Längenpolymorphismen in DNA-Bereichen, bei dem man folgende Schritte durchführt:

(a) Anlagerung von mindestens einem Primer-Paar an die zu analysierende DNA, wobei jeweils eines der Moleküle des Primer-Paares im wesentlichen komplementär ist zu einem der komplementären Stränge der 5'- bzw. 3'-Flanke einer simplen oder kryptisch simplen DNA-Sequenz, und die Anlagerung in solcher Orientierung erfolgt, daß die bei einer Primer-gesteuerten Polymerisationsreaktion mit jeweils einem der beiden Primer gewonnenen Syntheseprodukte nach einer Denaturierung als Matrize zur Anlagerung des jeweils anderen Primers dienen können.

Primer-gesteuerte Polymerasekettenreaktion
Abtrennung und Analyse der Polymerasekettenre-
aktionsprodukte

DE 3834636 A1

2007年12月15日，在“2007年中国年度人物”颁奖典礼上，
 中国首任驻美大使崔天凯、美国前驻华大使骆家辉、美国前
 国务卿奥尔布赖特、美国前副国务卿赖斯、美国前参议院
 外交关系委员会主席鲍威尔、美国前众议院外交关系委员会

Die simplen und kryptisch simplen DNA-Sequenzen
durften durch genomische Mechanismen entstanden
sein, die eine Tendenz haben, bereits existierende kurze
Verdopplungen von beliebigen DNA-Sequenzmotiven
nochmals zu verdoppeln oder in beliebigen DNA-Se-
quenzmotiven längere Bereiche von bereits bestehen-

Für das erfindungsgemäße Verfahren geeignete simple oder kryptisch simple DNA-Sequenzen können mit oder ohne Hilfe eines Computerprogramms in bereits bekannten DNA-Sequenzen gefunden werden (9). Ge-
wisse Sequenzen, wie zum Beispiel KSPACT, sind in DNA-Se-
quenzdatenbanken, wie zum Beispiel Genbank, bekannt. Sie
sowie den bereits und von Zufallssequenzen aus von
DNA-Sequenzen entnommene Wörter, sind für

kiert wird. Aus dem Bereich der flankierenden DNA-Sequenzen ohne interne Sequenzwiederholungen werden dann Stücke ausgewählt, zu denen geeignete komplementäre, synthetische Oligonucleotide hergestellt werden. Geeignet ist ein Oligonucleotid für diesen Zweck dann, wenn seine Nucleotidzusammensetzung und seine Nucleotidabfolge mit hoher Wahrscheinlichkeit nur einmal im zu untersuchenden Genom vorkommt und somit spezifisch für den individuell zu analysierenden DNA-Bereich ist.

Vorzugsweise werden im erfindungsgemäßen Verfahren Primer-Paare verwendet, deren einzelne Moleküle in einem Abstand von 50 bis 500 Nucleotiden voneinander an den zu untersuchenden DNA-Bereich angelagert werden, diesen also sozusagen im angegebenen Abstand umspannen. Dabei wird der zu untersuchende DNA-Bereich von den angelagerten Molekülen des Primer-Paares umgeben.

Vorzugsweise werden bei dem erfindungsgemäßen Verfahren zwei Primer-Paare eingesetzt. In einer besonders bevorzugten Ausführungsform werden 2 bis 50 Primer-Paare eingesetzt.

Vorzugsweise haben die beim erfindungsgemäßen Verfahren verwendeten Primer eine Länge von 15 bis 25 Nucleotiden.

In einer bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens werden bei der Verwendung mehrerer Primer-Paare die einzelnen Primer-Paare so ausgewählt, daß die entsprechenden spezifischen Polymerase-Kettenreaktionsprodukte der einzelnen Primer-Paare auf einem geeigneten Gel in einzelne Banden auftrennbar sind.

In einer anderen bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens erfolgt der Nachweis der spezifischen Polymerasekettenreaktionsprodukte durch radioaktive Markierung oder durch nicht-radioaktive Markierung, z.B. mit Fluoreszenzfarbstoffen.

Die Markierung der Oligonucleotid-Paare kann radioaktiv oder wie in (12) beschrieben, mit einem fluoreszierenden Farbstoff durchgeführt werden.

Ferner sind Bestecke (Kits), mit denen sich das erfindungsgemäße Verfahren durchführen läßt, Gegenstand der vorliegenden Erfindung. Die darin enthaltenen Primer sind gegebenenfalls radioaktiv, z.B. mit ^{35}S oder ^{14}C , oder fluoreszenzmarkiert.

Die beim erfindungsgemäßen Verfahren erhaltenen Syntheseprodukte lassen sich mit hochauflösenden Gel-systemen, wie üblichen Sequenzierungsgelen, auftrennen. Dabei kann auch die Länge der Syntheseprodukte bestimmt werden. Polymorphismen, die durch Insertionen oder Deletionen einzelner oder mehrerer Motive der simplen oder kryptisch simplen DNA-Sequenz gebildet werden, sind durch eine geänderte Position der Syntheseprodukte im Gel erkennbar. Bei geeigneter Auswahl der Primer-Paare lassen sich bei einem geeigneten A-Block gelssystem des Gelsystems mit 20 bis 50

zugrund der individuellen Kombination von Längenvorgängen der erhaltenen Syntheseprodukte zuverlässig feststellen.

Falls keine geeigneten simplen oder kryptisch simplen DNA-Sequenzen in den zu untersuchenden DNA-Bereichen bekannt sind, lassen sich diese wie folgt identifizieren.

Bei zu untersuchender genomischer DNA-Komplexen partiellen Restriktionsspaltung unterworfen. Dabei werden Restriktionsenzyme verwendet. Ein zu einer

weise nicht in simplen oder kryptisch simplen DNA-Sequenzen spalten. Die erhaltenen DNA-Fragmente werden in einem geeigneten Vektor, z.B. in lambda-Phagen-Derivaten oder in M13-Phagen cloniert und sodann in üblicher Weise durch Hybridisierung auf simple oder kryptisch simple DNA-Sequenzen abgesucht; vgl. (11). Die verwendeten Sondenmoleküle sind synthetische DNA-Moleküle, die verschiedene Permutationen von simplen oder kryptisch simplen DNA-Sequenzen enthalten. Auf diese Weise lassen sich hybridisierende Plaques identifizieren. Sodann kann die darin enthaltene rekombinante DNA isoliert und durch Sequenzierung charakterisiert werden. Die so erhaltene DNA-Sequenz läßt sich dann auf für das erfindungsgemäße Testverfahren geeignete DNA-Sequenzen absuchen.

Das erfindungsgemäße Verfahren wurde mit Drosophila-DNA als Modellsystem ausgeführt. Da jedoch simple und kryptische simple DNA-Sequenzen in allen eukaryontischen Genomen und teilweise auch in prokaryontischen Genomen vorkommen, ist davon auszugehen, daß die im Drosophila-Modellsystem erzielten Ergebnisse auch bei der Untersuchung anderer Genome, insbesondere bei der Untersuchung des menschlichen Genoms erzielt werden können.

Somit ist das erfindungsgemäße Verfahren für die Identitäts- und Verwandtschaftsbestimmung von Organismen, beispielsweise von Menschen geeignet.

Beim Menschen lassen sich Vaterschaftstests und forensische Tests zur Identitätsbestimmung von Straftätern mit dem erfindungsgemäßen Verfahren durchführen.

Neben der Identitätsbestimmung für Individuen ist das Verfahren auch geeignet, den Vererbungsgang für genetische Krankheiten zu bestimmen, für die der Locus bekannt und sequenziert ist. Dazu werden eine oder mehrere simple oder kryptisch simple Sequenzen ausgewählt, die im oder in der Nähe des zu analysierenden Locus liegen. Das spezifische Längenmuster dieser Bereiche wird mit dem mutierten Locus korreliert, so wie dies mit herkömmlichen RFLP-Markern üblich ist; vgl. (14). Mit dieser Information kann dann bei den betroffenen Familien eine genetische Beratung bzw. eine pränatale Diagnose durchgeführt werden und zwar in analoger Weise, wie es für RFLP-Marker üblich ist. Der Einsatz des erfindungsgemäßen Verfahrens zu diesem Zweck ist vor allem deswegen sinnvoll, weil es sich auf DNA-Bereiche stützt, die mit einer vorhersagbaren Wahrscheinlichkeit polymorph sind, während die RFLP-Analyse auf zufällig gefundene Variationen angewiesen ist, die oft weit von dem Locus selbst entfernt liegen, was die Diagnosesicherheit herabsetzt.

Ferner ist das erfindungsgemäße Verfahren zur Bestimmung von Polymorphismen in simplen oder kryptisch simplen DNA-Sequenzen von Tieren und Pflanzen geeignet. Daher können in der Tierzucht, z.B. bei Pferden, Hunden oder Rindern die Verwandtschaftsverhält-

nisse festgestellt werden. Das erfindungsgemäße Verfahren gegenüber den bisher bekannten Verfahren in seiner breiten Anwendbarkeit, schnellen Durchführbarkeit und in seiner hohen Empfindlichkeit. Der im erfindungsgemäßen Verfahren durchgeführte Amplifikationsschritt für die langenpolymorphen simplen oder kryptisch simplen DNA-Sequenzen macht einen unabhängigen Nachweisschritt, wie eine nachfolgende Hybridisierungsreaktion, überflüssig. Dadurch eignet sich das erfindungsgemäße Verfahren

besonders gut zur Automatisierung und für Routine-tests und Reihenuntersuchungen.

Die Figuren zeigen:

Fig. 1 Hybridisierung einer Genbank mit einer simplen DNA-Sequenz als Sondenmolekül.

Auf einer 12 x 12 cm großen Platte wurden ca. 20 000 vereinzelt Phagenclon ausplattiert und mit einem Sondenmolekül hybridisiert, das die simple Trinucleotidsequenz CAG/CTG enthält. Dabei werden etwa 300 bis 400 positive Signale erhalten. Die positiven Signale sind als Schwärzung erkennbar.

Fig. 2 Sequenz des in Beispiel 2 auf Polymorphismus getesteten Bereichs.

Die Bereiche, zu denen komplementäre Oligonucleotide synthetisiert wurden, sind mit einer Wellenlinie unterstrichen. Der Bereich der simplen DNA-Sequenz ist mit einer Doppellinie unterstrichen. Die direkte Wiederholung von 8 Nucleotiden ist mit zwei Pfeilen gekennzeichnet. Die HaeIII-Spaltstelle ist kursiv markiert.

Fig. 3 Analyse der Längenvariationen in 11 Wildtypstämmen von *Drosophila*.

Die mittels PCR amplifizierten und mit HaeIII gespaltenen DNA-Sequenzen sind in den Bahnen 1 bis 11 aufgetragen. Rechts ist eine Sequenzierungsreaktion aufgetragen, die als Längenmarker dient. Die Position der erwarteten Fragmente ist links mit Pfeilen markiert. Die Positionen der zusätzlich beobachteten Fragmentklassen ist mit Strichen markiert.

Fig. 4 Test auf Reproduzierbarkeit.

Zehn unabhängige PCR-Ansätze mit der DNA-Präparation "A" des *Drosophila*-Stammes Nummer 3 wurden links aufgetragen, zehn unabhängige PCR-Ansätze mit der DNA-Präparation "B" des *Drosophila*-Stammes Nr. 3 wurden rechts aufgetragen. Ganz rechts sind Markerfragmente aus einer Sequenzierungsreaktion aufgetragen. Alle beobachteten Testbanden sind identisch.

Die Beispiele erläutern die Erfindung.

Beispiel 1

Isolierung von Clonen, die simple DNA-Sequenzen enthalten.

Drosophila-DNA wird mit der Restriktionsendonuclease EcoRI vollständig gespalten und die resultierenden Fragmente werden in den λ -Vektor 641 cloniert. Eine nähere Beschreibung der angewendeten Methoden findet sich in (11). Somit wird eine Genbank erhalten, von der etwa 20 000 Phagen ausplattiert werden. Die entsprechenden vereinzelt Plaques werden auf einen Nitrocellulose-Filter übertragen und mit einem Sondenmolekül hybridisiert, das das simple DNA-Sequenzmotiv CAG/CTG enthält.

Die Filter werden bei 65°C hybridisiert und gewaschen. Die Hybridisierungslösung enthält 5x SSPE, 5x

SDS (die Zusammensetzung von 1x SSPE und 1x SDS (die Zusammensetzung von 1x SSPE und 1x SDS ist in (11) beschrieben).

Etwa 300 bis 400 der ausgebildeten Plaques zeigen ein positives Signal; vgl. Fig. 1. Von diesen Plaques werden einige gereinigt, es wird DNA isoliert und sequenziert. In den erhaltenen DNA-Sequenzen lassen sich Bereiche identifizieren, die die simple DNA-Sequenz CAG/CTG

Beispiel 2

Nachweis von Längenpolymorphismen

Für diesen Versuch wurde die in Fig. 2 wiedergegebene und in (13) veröffentlichte DNA-Sequenz ausgewählt. Es wurden zwei Oligonucleotide mit den folgenden Sequenzen synthetisiert:

Oligonucleotid 1: 5'-TAAGCTTGGGAATCA-3'
Oligonucleotid 2: 5'-ATTGAACCTTGTATC-3'.

Diese DNA-Sequenzen befinden sich unmittelbar am Beginn bzw. am Ende der in Fig. 2 dargestellten Sequenz. Zur Verwendung als Primer werden die synthetisierten Oligonucleotide an ihrem 5'-Ende mit ^{32}P markiert. Sodann wird eine PCR-Reaktion mit den markierten Primern durchgeführt. Insgesamt werden 20 Zyklen durchgeführt, wobei jeweils 90 Sekunden bei 95°C denaturiert wird, 90 Sekunden bei 45°C angelagert wird und 120 Sekunden bei 72°C synthetisiert wird. Als zu untersuchende DNA's werden die genomischen DNA's von 11 Wildtypstämmen von *Drosophila melanogaster* aus verschiedenen Gebieten der ganzen Welt eingesetzt. Diese *Drosophila*-Wildtypstämme stammen ursprünglich von einzelnen fertilisierten Weibchen ab und wurden während der letzten 10 Jahre gesammelt. Nach der PCR-Reaktion werden die amplifizierten Fragmente mit der Restriktionsendonuclease HaeIII gespalten. Dabei sollten üblicherweise zwei Fragmente entstehen, die eine Länge von 202 Nucleotiden bzw. 177 Nucleotiden aufweisen. Dieser Schritt ist für Routineexperimente normalerweise nicht erforderlich. Er dient hier lediglich der Verfeinerung der Analyse. Die entstehenden Fragmente werden auf einem 5prozentigen Sequenzierungsgel aufgetrennt, sodann wird das Gel getrocknet und ein Röntgenfilm wird mit dem getrockneten Gel belichtet. Die beiden erwarteten DNA-Fragmente zeigen einen ausgeprägten Polymorphismus in den verschiedenen *Drosophila*-Wildtypstämmen. Das die simple DNA-Sequenz enthaltende Fragment mit 202 Nucleotiden zeigt vier verschiedene Größenklassen; vgl. Fig. 3. Diese Größenklassen sind jeweils um drei Nucleotide verschoben. Dies ist ausgehend von Rasterverschiebungen innerhalb der Wiederholung der Trinucleotide zu erwarten. In drei Fällen tauchen gleichzeitig zwei verschiedene Banden auf; vgl. Fig. 3, Bahnen 5, 8 und 9. Dies ist dadurch erklärbar, daß in diploiden Organismen jeder Locus zweimal vorkommt und mit unterschiedlichen Allelen besetzt sein kann (sogenannter balancierter Polymorphismus). Die Bande des Fragments mit 177 Nucleotiden zeigt drei verschiedene Größenklassen, die 5 bzw. 8 Nucleotide auseinanderliegen; vgl. Fig. 3. Die um 8 Nucleotide kürzere Bande ist vermutlich durch eine Deletion der in der DNA-Sequenz markierten Wiederholung von 8 Nucleotiden entstanden.

Die Ergebnisse der Analyse sind in Fig. 3 dargestellt. Die Banden sind in der Abbildung mit Pfeilen markiert. Die Banden, die nicht den erwarteten Fragmenten entsprechen, sind mit Strichen markiert.

Die Mehrzahl der in diesem einfachen Experiment untersuchten Stämme ist bereits voneinander unterscheidbar. Nicht unterscheidbar sind lediglich die Stämme 2, 7 und 11 sowie 3 und 4. Für einen tatsächlichen Test wurde man daher weitere Primer-Paare einsetzen. Beispielsweise könnten 20 bis 50 unabhängige DNA-Bereiche untersucht werden, um eine eindeutige Identifizierung zu ermöglichen. Da die Größenklassen der ein-

zelen Drosophila-Wildtypstämme in sich einheitlich sind, ist davon auszugehen, daß die beobachteten Polymorphismen nicht mit so hoher Häufigkeit auftauchen, daß Verwandtschaftsbeziehungen nicht mehr feststellbar wären. Die Drosophila-Wildtypstämme stammen ja alle von jeweils einem einzelnen Ausgangspaar ab und für den Test wurde die DNA von mehreren 100 Individuen vereinigt. Wenn innerhalb tiefer "Familien" eine Änderung des Musters aufgetreten wäre, müßte man mehr als maximal zwei Banden erwarten. Dies ist aber hier nicht der Fall. Daraus folgt, daß die beobachteten Längensklassen zumindest für einige Dutzend Generationen stabil sind.

Beispiel 3

Test auf Reproduzierbarkeit

Die beobachteten Längenvariationen könnten auch durch während des Versuchs auftretende Polymerasefehler verursacht werden. Um diese Möglichkeit auszuschließen und um gleichzeitig die allgemeine Reproduzierbarkeit nachzuweisen, wird der in Beispiel 2 durchgeführte Versuch mit zwei verschiedenen DNA-Präparationen des Drosophila-Stammes Nr. 3 in jeweils 10 unabhängigen Ansätzen wiederholt. Aus Fig. 4 ist ersichtlich, daß alle Ansätze zu den gleichen Banden führen. Ähnliche Versuche wurden auch für andere Loci durchgeführt. Jedoch wurde in keinem Fall eine Veränderung der Bandenlänge beobachtet. Dies zeigt, daß das Verfahren zuverlässig reproduzierbar ist.

Literatur:

1. EP-A2 02 00 362
2. R. K. Saiki et al., Science 239 (1988), 487—491
3. A. J. Jeffreys et al., Nature 314 (1985), 67—73
4. A. J. Jeffreys et al., Nature 316 (1985), 76—79
5. EP 87 11 6408.3
6. D. Tautz, Doktorarbeit Universität Tübingen (1983)
7. D. Tautz und M. Renz, J. Mol. Biol. 172 (1984), 229—235
8. D. Tautz und M. Renz, Nucleic Acids Research 12 (1984), 4127—4138
9. D. Tautz et al., Nature 322 (1986), 652—656
10. G. Levinson und G. A. Gutman, Mol. Biol. and Evolution 4 (1987), 203—221
11. T. Maniatis, E. F. Fritsch und J. Sambrook, "Molecular Cloning, a Laboratory Manual", Cold Spring Harbor Laboratory, New York, 1982
12. L. M. Smith et al., Nature 321 (1986), 674—679
13. K. A. Wharton et al., Cell 40 (1985), 55—62
14. Y. W. Kan und A. M. Dozy, Proc. Natl. Acad. Sci. USA 75 (1978), 5631—5635

steuerten Polymerisationsreaktion mit jeweils einem der beiden Primer gewonnenen Syntheseprodukte nach einer Denaturierung als Matrize zur Anlagerung des jeweils anderen Primers dienen können;

(b) Primer-gesteuerte Polymerasekettenreaktion; und

(c) Auftrennung und Analyse der Polymerasekettenreaktionsprodukte.

2. Verfahren nach Anspruch 1, bei dem die Moleküle des Primer-Paares in einem Abstand von 50 bis 500 Nucleotiden voneinander an die zu untersuchende DNA angelagert werden.

3. Verfahren nach Anspruch 1 oder 2, bei dem mindestens 2 Primer-Paare eingesetzt werden.

4. Verfahren nach Anspruch 1 oder 2, bei dem 2 bis 50 Primer-Paare eingesetzt werden.

5. Verfahren nach einem der Ansprüche 1 bis 4, bei dem die Primer eine Länge von 15 bis 25 Nucleotiden aufweisen.

6. Verfahren nach einem der Ansprüche 1 bis 5, bei dem die Lage der Primer-Paare so ausgewählt wird, daß die spezifischen Polymerasekettenreaktionsprodukte der Primer-Paare auf einem geeigneten Gel in einzelne Banden auftrennbar sind.

7. Verfahren nach einem der Ansprüche 1 bis 6, bei dem der Nachweis der spezifischen Polymerasekettenreaktionsprodukte durch radioaktive Markierung oder durch nicht-radioaktive Markierung, wie Fluoreszenzfarbstoffe, erfolgt.

8. Verfahren nach einem der Ansprüche 1 bis 7, bei dem die zu untersuchenden simplen oder kryptisch simplen DNA-Sequenzen in der Nähe oder innerhalb eines genetisch definierten Locus sitzen, so daß der gefundene Polymorphismus als Marker für diesen Locus dienen kann.

9. Besteck (Kit) zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 8, enthaltend:

(a) ein oder mehrere Gefäße mit einer äquimolaren Mischung von 1 bis 50 Oligonucleotidprimer-Paaren, die simple oder kryptisch simple DNA-Sequenzen flankieren, wobei die Primer gegebenenfalls radioaktiv oder fluoreszenzmarkiert sind;

(b) ein Gefäß mit einem Enzym zur Polymerisation;

(c) ein Gefäß mit den vier Desoxynucleosidtriphosphaten; und

(d) ein Gefäß mit einer geeigneten Pufferstammlösung; und gegebenenfalls

(e) ein Gefäß mit einer Kontroll-DNA, die zum Testen der Komponenten des Bestecks (Kit) geeignet sind.

Hierzu 4 Seite(n) Zeichnungen

ZEICHNUNGEN SEITE 1

Nummer:

Int. Cl.⁸:

Offenlegungstag:

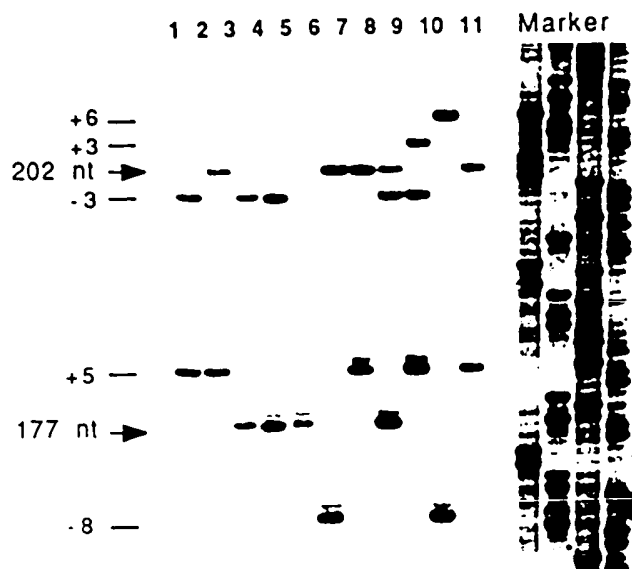
DE 38 34 636 A1

C 12 Q 1/68

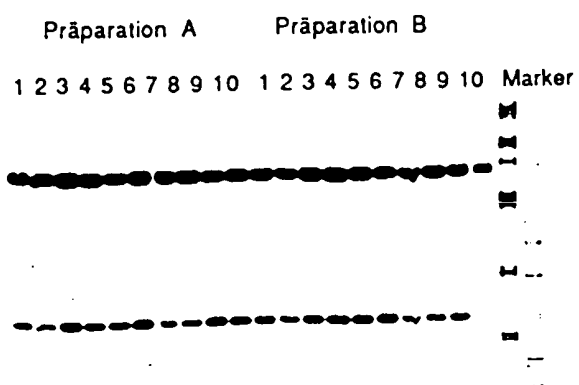
19. April 1990



TAAGCTTGGG AATCA TCTOG OOGAOGGGCA GOGATATGGG CATCATGCTC G0000G000C
 AATCCTOGAA GAATAGTGCA ATAATGCAAA CGATATCAOC CCAGCAACAG CAGCAGCAGC
AGCAGCAGCA AC-AGCAGCAA CATCAGCAGC AGCAACAGCA GCAGCAACAG CAGCAGCAGC
 202 nt | 177 nt
AACAGCAGCA GCAA CTCGGA GGOCTGGAGT TOGGTTCAGA GGGCTTGGAC CTGAATGGAT
 Hae III
 TTTGTGGATC TCOGGGTAAG TGGTCACTCA TGATGGACTC TATGGACTCG CTAACTAGCT
AACTAATCAT TCTACCATCC CAACTTGCAG ACTCATTTCA CTGGGTCAA ATGAATCCGC
 CCTCGATACA AAGTTCAAT



Figur 3



Figur 4

